Biosystems Engineering (2005) **91** (4), 513–524
doi:10.1016/j.biosystemseng.2005.04.003
AP—Animal Production Technology

Available online at www.sciencedirect.com
SCIENCE DIRECT

Evaluating Modelling Techniques for Cattle Heat Stress Prediction

T.M. Brown-Brandl[1]; D.D. Jones[2]; W.E. Woldt[2]

[1]USDA-ARS US Meat Animal Research Center, P.O. Box 166, Clay Center, Nebraska USA; e-mail of corresponding author: brandl@email.marc.usda.gov
[2]University of Nebraska, Lincoln, NE, USA

(Received 13 September 2004; accepted in revised form 12 April 2005; published online 7 July 2005)

Researchers have traditionally predicted animal responses by means of statistical models. This study was conducted to evaluate modelling techniques. One hundred and twenty-eight feedlot heifers were observed during a 2-month period during the summer of 2002. Respiration rate and surface temperature were taken on a random sample of 40 animals twice a day. Five different models (two statistical models, two fuzzy inference systems, and one neural network) were developed using 70% of this data, and then tested using the remaining 30%. Results showed that the neural network described the most variation in test data (68%), followed by the data-dependent fuzzy model (Sugeno type) (66%), regression models (59 and 62%), while the data-free fuzzy model (Mamdami type) described only 27%. While the neural-network model may be a slightly better approach, the researcher may learn more about responses using a fuzzy inference system approach. For all models tested, respiration rate is over-predicted at low stress conditions and under-predicted at high stress conditions. This suggests that all models are lacking a key piece of input data, possibly the accumulative effects of prior weather conditions, to make an accurate prediction.
Published by Elsevier Ltd on behalf of Silsoe Research Institute

1. Introduction

Heat stress in cattle causes decreases in feed intake and growth, and in extreme cases can cause death, resulting in substantial lost revenue to producers. An extreme example is a heat wave that occurred in western Iowa, July 1995, causing an approximate loss of 3750 head; direct losses were estimated at US \$2·8 million, and production losses at US \$28 million (Busby & Loy, 1996).

A method of predicting stressful environments would aid the producer in taking proactive steps in reducing stress. In cattle, respiration rate has been shown to increase with increasing temperature (Hahn *et al.*, 1997). Although the relationship between respiration rate and dry-bulb temperature is not a linear relationship, it has several advantages over other physiological measurements that make it an attractive alternative. The first advantage is that no special equipment or extensive training is needed to measure respiration. The second advantage is measurements of respiration rate can be done at some distance. As the animal is not moved from the pen, or confined in any way, measurement of respiration rate causes little or no additional stress on the animal. With these advantages, it appears that respiration rate could be an effective way for a producer to monitor stress.

Eigenberg *et al.* (2005) developed a Livestock Safety Monitor for cattle using a commercial weather station, and a small companion unit which calculates and displays danger level based on predicted respiration rate. Currently, the most challenging component in the system seems to be the prediction equation. If a producer is going to rely on the system to give warnings on stress level, it needs to be fairly accurate, especially under stressful conditions. If the prediction equation under-predicts respiration rate during the stressful events, the producer might not take appropriate action, while if the equation over-predicts the situation, the producer may begin to ignore warnings. The equation used (Eigenberg *et al.*, 2005) was developed with a standard multiple linear regression approach. The fit of

 Published by Elsevier Ltd on behalf of Silsoe Research Institute

that model to an independent dataset was not conclusive. The equation only accounted for 56% of total variation. Other modelling methods of predicting stress level may provide predictions that are more accurate.

Both regression models and soft computing techniques have been successfully used in predicting many biological parameters. The soft computing techniques implied here include fuzzy inference systems and neural networks. The use of soft computing techniques is a means of modelling that allows apt descriptions of mapping inputs to outputs.

All modelling schemes, whether based on traditional mathematical principles or developed through soft computing techniques, represent mapping a set of inputs to a set of outputs. Many models are developed without complete knowledge of the system being interpreted or predicted. For instance, analytical models are usually satisfactory at predicting outputs, but usually oversimplify the system.

Independent of the method used, all traditional types of models impose a form of mapping based on known information. A set of conventions used to create a form or outline must be assumed in order to develop the model. An alternative method is to use a data-free form to map a set of inputs to an output. In this case, natural rules are developed from data rather than imposing rules on the modelling system. Therefore, in a data-free system, rules are developed through clustering algorithms that divide data into natural partitions. Mapping is then optimised through various techniques. Therefore, the result of this data-free model is still mapping the inputs to outputs, similar to traditional algorithms. The rules represent mapping, whether imposed by the modeller or determined from data, and can be crisp (either true or false) or fuzzy (a continuum along the unit interval).

Applications of soft computing techniques to model complex physical and biological systems are well documented; following are examples—transport of various compounds through skin (Pannier *et al.*, 2003); detection of neonate behaviour (Belal *et al.*, 2002); nutrient flow in a constructed wetland for wastewater treatment (Woldt *et al.*, 2002). Modelling of complex systems involving multi-scale environmental processes has been described by Woldt *et al.* (2003), in which life cycle assessment and production systems are simulated in a framework of imprecision. Ganjyal *et al.* (2003) modelled properties of extruded starch using neural network. Thai and Shewfelt (1991) compared use of neural network and statistical methods to model sensory colour quality of produce.

2. Objectives

Objectives for this experiment were to develop and evaluate five types of models: two statistically based regression models, two types of fuzzy inference systems, and a neural network to predict respiration rate of feedlot cattle based on environmental variables and an indicator of breed.

3. Materials and methods

One hundred twenty-eight feedlot heifers of four breeds—(Angus, Charolais, Gelbvieh, and MARC III cross-bred [Pinzgauer, Red Poll, Hereford, Angus]) from the USDA-ARS Meat Animal Research Center (MARC) population were selected for this study. Generally, Angus cattle are black, MARC III are dark red, Gelbvieh are tan, and Charolais are white. Heifers initially weighing $384{\cdot}9 \pm 46{\cdot}5$ kg were assigned to one of four adjacent pens by breed (32 heifers/pen). Heifers were implanted with Synovex-H (200 mg testosterone propionate and 20 mg esteordial benzoate) before the study began to simulate standard feedlot practice. Heifers were fed twice daily, before 0800 h and after 1300 h, and had continual access to water.

Throughout the study, weather data (dry-bulb temperature t_{db} in °C, dew point temperature t_{dp} in °C, solar radiation r_s in W/m^2, wind speed v_w in m/s, and wind direction) were collected using an automated weather station (Vantage Pro, Davis Industries) located in the middle of the set of four pens. Ten days prior to initiating experiments, cattle were pre-conditioned to observers. During the pre-conditioning period, two observers spent 1 h twice daily walking outside pens.

Measurements of respiration rate and surface temperature were made twice daily (0800 and 1430 h) on a pre-determined schedule for a total of 30 days between mid-June and mid-August. On scheduled experimental days, two observers, working independently, randomly selected five animals per pen to observe. For each selected animal, identification number and respiration rate were recorded. Respiration rates were determined by visual observation of flank movement, and timing 10 breaths with a stopwatch. After respiration rate measurements on heifers were completed, surface temperature measurements were taken using an infrared thermometer (Raynger ST80 ProPlus; ± 1 °C) on the animals' back at approximately the last rib at a distance of less than 3 m. Surface temperature measurements were recorded on five heifers per pen.

3.1. *Modelling*

Data from any animal (16 of 128) that had been treated for pneumonia in the past were eliminated from the dataset. The dataset was then randomly divided into two subsets. Set 1 consisted of 1400 independent data points, and was used in the model development. Set 2 consisted of the remaining 600 data points, and the model evaluation was completed using this set. This ratio of testing and evaluation subsets was selected, as it is a common ratio to use when modelling. To quantify breed, average afternoon hair coat surface temperatures t_s in °C were used in the models (Angus—42·7; MARC III—42·1; Gelbvieh—40·8; Charolais—38·4) (Brown-Brandl *et al.*, 2003). Slope, intercept mean error, coefficient of determination, and a plot of residuals were recorded for each model.

3.1.1. *Regression models*

Two multiple regression models were developed using the regression procedure in SAS (SAS, 1999). The first model was a linear regression model and included breed, and the four weather parameters (t_{db}, t_{dp}, r_s, and v_w). The second model was a quadratic model and included breed, the four weather parameters (t_{db}, t_{dp}, r_s, and v_w), and interaction and quadratic effects of weather parameters. Parameters were considered significant if the probability was less than 0·05. Per cent of contribution was determined by calculating the ratio of each standardised regression estimate, squared by the sum of squares from all standardised regression estimates.

3.1.2. *Data-dependent fuzzy inference system*

The data-dependent model consisting of five inputs (breed, t_{db}, t_{dp}, r_s, and v_w) and one output (respiration rate), was implemented using Matlab®, (MathWorks, Inc., 2000), a Sugeno type of approach (Takagi & Sugeno, 1985). This fuzzy inference system was developed using the aforementioned 'dataset 1' as a training set, and 'dataset 2' as the testing set. The inference system was developed using 'genfis2' (which utilises subtractive clustering) within Matlab®, and the 'ANFIS' neural-network training routine. This approach is described in more detail in Appendix A.

3.1.3. *Data-free fuzzy inference system*

Developing a Mamdani type fuzzy inference system (Mamdani & Assilian, 1975) begins in a manner similar to the Sugeno type fuzzy inference system. However, the essence of this approach requires a great deal more human intervention on the part of the modeller. In fact, it is possible to build a Mamdani type fuzzy inference system based completely on heuristics and conventional wisdom, which results in a model that is data free. In this case, this model was implemented in FuzzyTech (Inform, 1999). While it is possible to construct a data-free model, the first step in Mamdani model development, available data can help to refine the model. Data clustering was completed using the Fuzzy c-Means method. The Mamdani modelling approach does not lend itself to being trained in a manner similar to 'ANFIS' method in Mathlab® because the consequence is not differentiable. Instead, FuzzyTech implements a slightly different approach to neural-network-based tuning of the rule system. In this case, original membership functions are retained, and degree of support for each rule is determined such that error between known data and simulated data is minimised. A fuzzy inference system was developed using the previously described 'dataset 1' for training, and 'dataset 2' for testing the resultant model. More details on this method can be found in Appendix B.

3.1.4. *Neural-network model*

A neural-network model was developed using the aforementioned data arrangement. The development system used in this work was a system that is commercially available (NeuroShell 2·0, Ward Systems Group Inc., 1995, Frederick, MD). Modelling parameters included number of hidden layers, number of nodes in hidden layer(s), learning rate, momentum, and weights. Refinement of parameters was sequentially adjusted in the aforementioned order such that error was minimised for each parameter. To avoid over fitting of the model more aggressive refinement was not attempted. The resulting neural-network model acts as a predictor of respiration rate for observations of breed, t_{db}, t_{dp}, r_s, and v_w. See Appendix C for a more details on this method.

4. Results

4.1. *Regression models*

The linear regression model for respiration rate R_R in breaths/min was developed using only linear effects of breed using average afternoon hair coat surface temperatures t_s in °C (Angus—42·7 °C; MARC III—42·1 °C; Gelbvieh—40·8 °C; Charolais—38·4 °C; Brown-Brandl *et al.*, 2003), and four weather parameters. All parameters except t_{dp} were significant; however, t_{dp} was left in the model for completeness [Eqn (1)]

$$R_R = -183{\cdot}6 + 4{\cdot}22t_s + 3{\cdot}89t_{db} - 0{\cdot}07t_{dp} - 1{\cdot}53v_w + 0{\cdot}03r_s \tag{1}$$

The equation accounts for 60·4% of total variation in data. By far the largest portion of variation is accounted for by t_{db} (68·7%), and then v_w (13·3%), r_s (9·2%), breed (8·8%), and finally t_{dp} for almost none of the variation.

When this model was tested against the remaining 30% of data, the model was able to account for 59% of total variation in respiration rate (Table 1). Slope was equal to 0·62, and intercept was 35·8. Plot of residuals appeared to have a slight quadratic pattern [*Fig. 1*(*a*)]. It has been well documented that respiration rate has a non-linear response to t_{db} (Hahn *et al.*, 1997). Because this model does not account for that non-linear response, this could result in the pattern found in the residuals.

Table 1
Statistical results of each of four models when compared to 600 point test dataset

Model	*Slope*	*Intercept*	R^2	*Mean error*
Linear regression	0·62	35·8	0·59	1·14
Quadratic regression	0·65	32·5	0·62	0·91
Data-dependent fuzzy model	0·70	28·0	0·66	0·92
Data-free fuzzy model	0·31	70·0	0·27	8·0
Neural network	0·70	26·3	0·68	1·04

R^2, coefficient of determination.

The quadratic regression equation was developed in two stages. The model was first run with all main effects, quadratic effects, and interaction effects. Then a final model was developed using all significant effects found in the first run on the model. The final model contains effects of breed, t_{db}, t_{dp}, v_w, r_s, t_{db}^2, t_{dp}^2, v_w^2, $t_{db} \times t_{dp}$, and $t_{db}v_w$ [Eqn (2)]:

$$\begin{aligned} R_R = & -223{\cdot}7 + 4{\cdot}23t_s - 4{\cdot}10t_{db} + 12{\cdot}42t_{dp} + 3{\cdot}50v_w \\ & + 0{\cdot}03r_s + 0{\cdot}12t_{db}^2 - 0{\cdot}50t_{dp}^2 + 0{\cdot}3v_w^2 \\ & + 0{\cdot}22t_{db}t_{dp} - 0{\cdot}21t_{db}v_w \end{aligned} \quad (2)$$

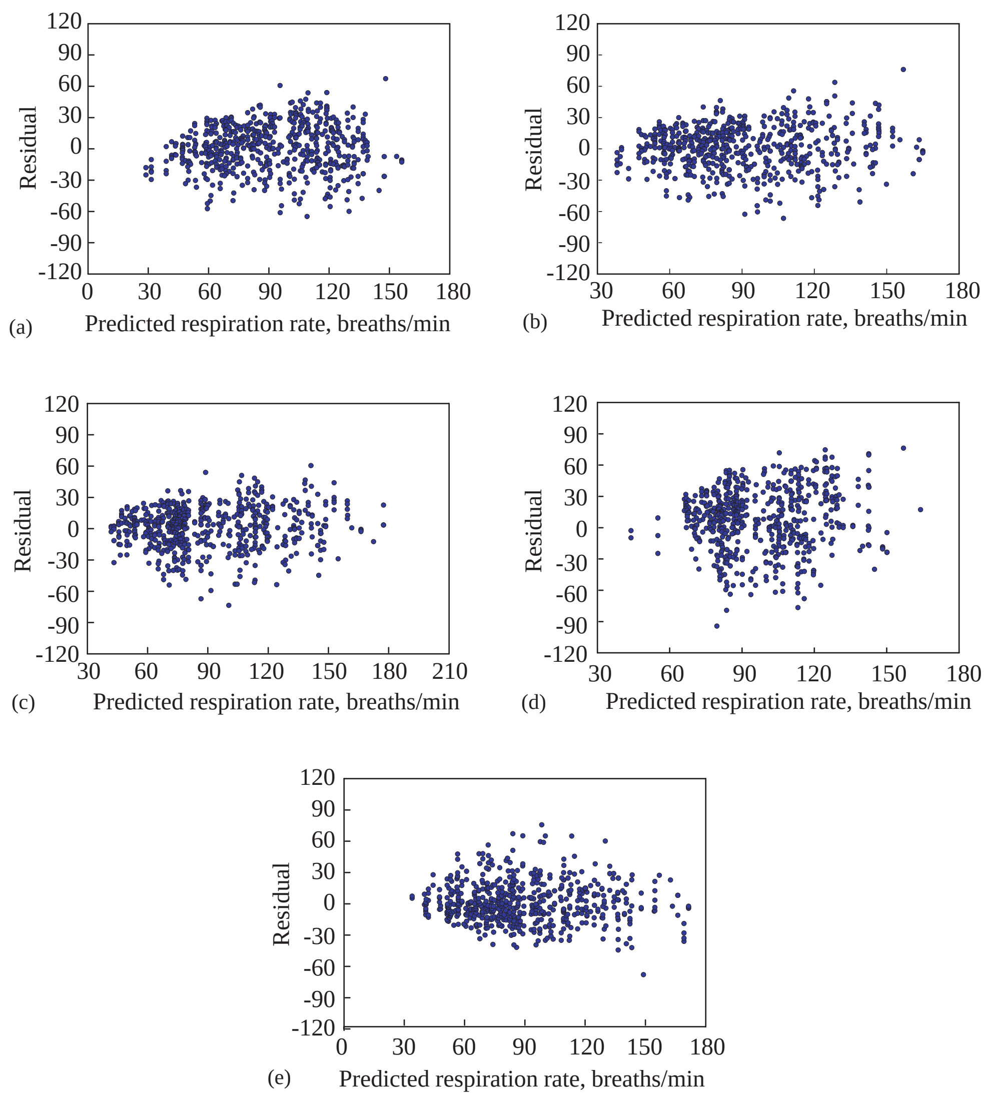

Fig. 1. Residual plots: (a) for linear regression model; (b) for quadratic regression model containing linear, quadratic, and interaction terms; (c) for data-dependent type fuzzy inference system; (d) for data-free type fuzzy system; (e) for the neural-network model

Table 2
Distribution of variation associated with each of parameters in the quadratic regression model, as determined by the ratio of standardised regression estimates

Parameter	*Total variation explained, %*
Main effects	26·3
Cattle breed	0·8
Dry-bulb temperature t_{db}, °C	6·7
Dew-point temperature t_{dp}, °C	11·6
Wind speed v_w, m/s	6·1
Solar radiation r_s, W/m^2	1·1
Interaction effects	31·7
$t_{db} \times t_{dp}$	9·8
$t_{db} \times v_w$	21·9
Quadratic effects	42·0
t_{db}^2	17·4
t_{dp}^2	24·1
v_w^2	0·5

The addition of five terms only accounted for an additional 4·0% of variation, for a total of 64·4%; however, distribution of variation was changed considerably. Main effects only accounted for 26·3%, quadratic effects 42·0%, and interaction effect the remaining 31·7% (Table 2).

When this model was tested against the remaining 30% of data, the model was able to account for over 62% of total variation in respiration rate (Table 1). The slope was equal to 0·65, and intercept was 32·5. The plot of residuals appears mostly random [*Fig. 1*(*b*)]. However, there appears to be less variation in errors at very low respiration rate, possibly due to low number of points.

4.2. *Data-dependent fuzzy inference system*

The developed model was a set of seven rules dictating mapping of inputs to output. Shown in *Fig. 2* is an example of the interactive interface that describes the data-dependent fuzzy inference system. Each row in the figure represents one rule, and consists of five membership functions corresponding to each of the five inputs (breed, t_{db}, t_{dp}, r_s, and v_w). The membership functions are shown in the first five columns. Position of the solid vertical line represents input value. In this example, breed was set at 40·8 (Gelbvieh), and the data-assigned values were t_{db} of 26·1 °C, r_s of 460 W/m^2, v_w of 16·4 m/s, and t_d of 18·5 °C. For each individual membership function, range of input values is represented by values on the X axis and membership value represented on the Y axis. The shaded region is a visual representation of the resulting membership of input value. In Rule 2, for example, the shaded region of the breed membership function is very small, therefore, this rule applies approximately 5% of this input. In Rule 5, the shaded region of the breed membership function is about 80%, showing this rule is very applicable to this input. The last column represents output (respiration rate) of each of the seven rules. Position of the column represents the output value (respiration rate). The portion of the bar that is black represents the weighting factor for that rule, and is determined by the minimum membership value in each rule—the horizontal line and arrow shows which input function determines the weighting factor. A single output is a result of a weighted average of output from each of the seven rules, and is shown in the lower right hand corner. Some rules (*e.g.* Rule 6 in this example) contribute little to output, while others contribute significantly (*e.g.* Rule 5 in this example).

In modelling terms, inputs are combined as fuzzy clusters described with Gaussian membership functions. The model uses a logical 'AND' relationship to combine data space into fuzzy clusters. The degree of belonging of an input vector to a particular cluster defines contribution of the associated rules. Ultimate output is a weighted average of each contributing rule.

Examining rules presented in *Fig. 2* reveals several interesting features about the relationship between respiration rate and inputs. There is a variety of means to quantitatively compare impact of each rule and variable on output. However, for this discussion, those assessments will remain qualitative.

Consider the second column of *Fig. 2*, which represents t_{db}. Projections of each cluster on the 't_{db}' dimension are spaced across the data dimension with marginal similarity between rules. This indicates that respiration rate is sensitive to t_{db}.

In contrast, consider the fourth column of *Fig. 2*, which represents v_w. Projections of each cluster on 'v_w' dimension are spatially very similar regardless of rule. The obvious exception is the projection associated with Rule 6, which shows some deviation from other clusters. Nonetheless, it can be argued that v_w has little influence on respiration rate.

Also of interest is the first column of *Fig. 2*, which represents breed. Projections associated with Rules 1, 2, 5, and 6 tend to be very similar to each other. Shapes of Gaussian projections are different, yet they have similar centroids and represent approximately the same portion of 'breed' dimension. Projections associated with Rules 3 and 4 also tend to be very similar to each other. One could assert that there are two groupings of breed that impact respiration rate given weather conditions experienced. Projection associated with Rule 7 tends to span the entire 'breed' dimension, indicating that (for

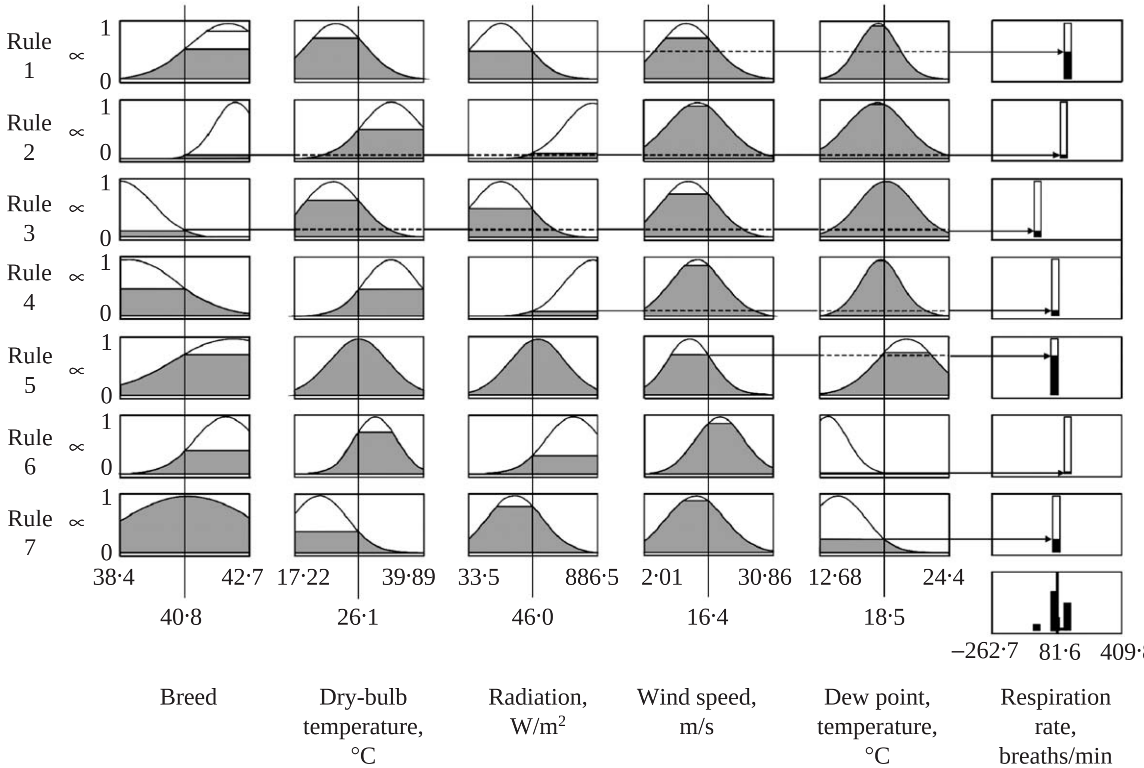

Fig. 2. An example of the interactive interface generated by Matlab (Mathworks, Inc., 2000) describing the data-dependent fuzzy inference system. In this example, the inputs as illustrated above were breed of 40·8 (which represents Gelbvieh), dry-bulb temperature of 26·1 °C, solar radiation of 460 W/m^2, wind speed of 16·4 m/s, dew point temperature of 18·5 °C. The resulting output from each of the seven rules is illustrated by the position of the vertical bar in the column labelled respiration rate. The weighting factor determined by the minimum degree of support μ of the five parameters, for each of the seven outputs is illustrated by the shaded region of the bar. The last graph in this column shows the relative contributions of the seven rules to the final output

this rule) breed is not important. Closer examination of weather conditions associated with this rule ('low' t_{db}, 'low' r_s, 'low' v_w, and 'low' t_{dp}) indicates that is a logical assertion.

Further examination and parsing of individual rules and collection of rules is warranted and could continue. However, performance of the model is described by comparing predicted and observed respiration rate. When this model was tested against the remaining 30% of data, the model was able to account for over 66% of total variation in respiration rate (Table 1). The slope was equal to 0·70 and intercept was 28·0. The plot of residuals appears mostly random [*Fig. 1(c)*].

4.3. *Data-free fuzzy inference system*

The data-free fuzzy inference system that results from application of the Mamdani approach is different than previously described data-dependent fuzzy inference system. In this case, the inference system can be described with a set of membership functions (*Fig. 3*) that were constructed, based on linguistic descriptors of input conditions. For example, r_s at levels below 250 W/m^2 are considered cloudy conditions, r_s levels from 250 to 400 W/m^2 have decreasing membership in the 'cloudy' set, and r_s levels above 400 W/m^2 are not cloudy. This linguistic expression of environmental conditions and breed represents a model component that is dependent on human interpretation of data.

The fuzzy inference system also consists of rules that process input data in conjunction with membership functions. In this case the rule set consists of 20 rules is shown in Table 3. The final rule set is reduced from a complete (*i.e.* total enumeration) set of rules that represent all combinations of input membership functions. Thus, the complete rule set for this inference system began with 960 rules. Clustering of dataset 1 was completed to reduce training complexity, and a set of 53 clusters were identified from the 1400 original data. Neural-network training of the complete set of rules, using the 53 clusters, resulted in 20 rules that describe

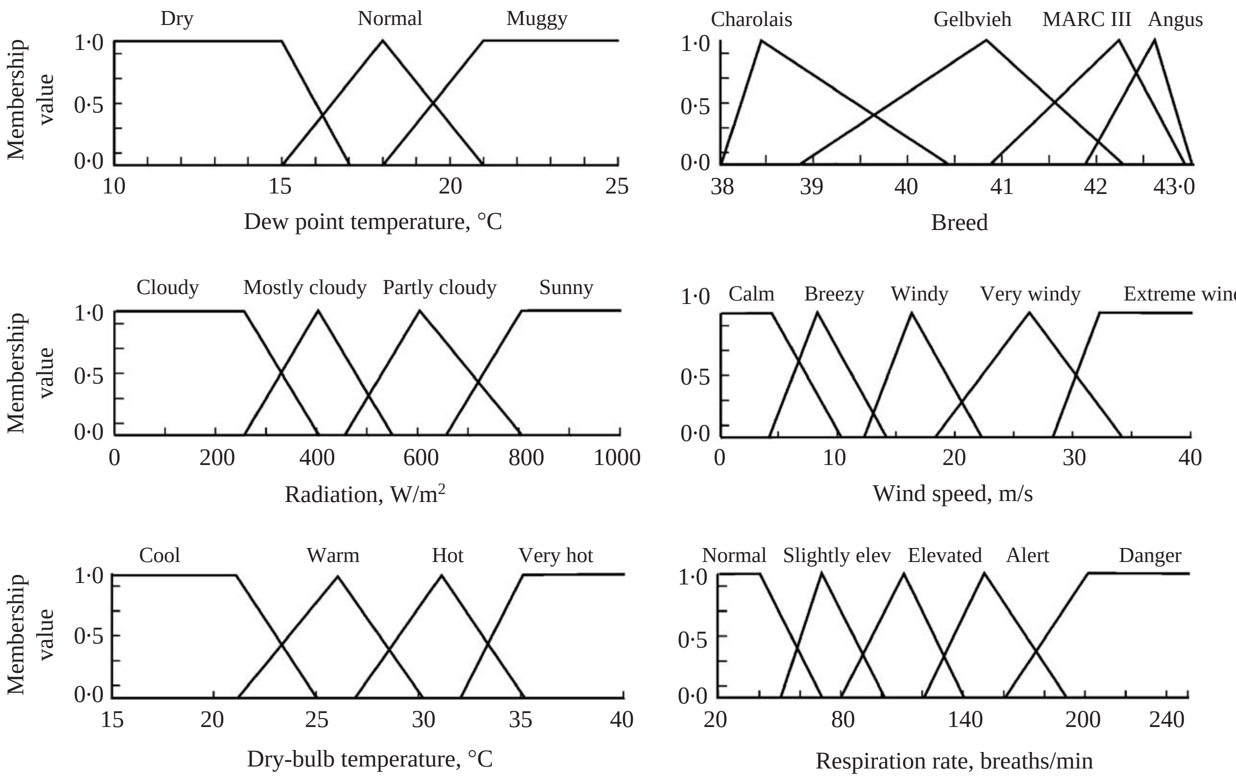

Fig. 3. The membership function structure developed for the data-free fuzzy inference system using FuzzyTech

Table 3

Rule structure for data-free fuzzy inference system (Mamdani) developed using the FuzzyTech (by Inform). If the set of conditions are used as inputs (first five columns), then the model predicts the respiration rate level (last column)

If… Dew point condition	*Cattle breed*	*Solar radiation condition*	*Dry-bulb temperature condition*	*Wind Condition*	*Then… Respiration rate level*
Dry	Charolais	Cloudy	Cool	Windy	Normal
Dry	Charolais	Sunny	Hot	Breezy	Alert
Dry	Gelbvieh	Partly cloudy	Cool	Windy	Normal
Dry	Gelbvieh	Sunny	Cool	Breezy	Normal
Dry	MARC III	Mostly cloudy	Cool	Breezy	Normal
Dry	MARC III	Partly cloudy	Very hot	Windy	Alert
Dry	Angus	Partly cloudy	Very hot	Calm	Alert
Normal	Charolais	Partly cloudy	Warm	Windy	Slightly elevated
Normal	Gelbvieh	Cloudy	Cool	Calm	Normal
Normal	Gelbvieh	Mostly cloudy	Cool	Breezy	Normal
Normal	Gelbvieh	Sunny	Cool	Windy	Normal
Normal	Gelbvieh	Sunny	Warm	Breezy	Slightly elevated
Normal	Angus	Cloudy	Cool	Calm	Normal
Muggy	Charolais	Sunny	Very hot	Windy	Danger
Muggy	Gelbvieh	Mostly cloudy	Very hot	Calm	Danger
Muggy	Gelbvieh	Sunny	Warm	Very windy	Slightly elevated
Muggy	MARC III	Cloudy	Warm	Windy	Slightly elevated
Muggy	MARC III	Sunny	Hot	Calm	Danger
Muggy	Angus	Mostly cloudy	Cool	Windy	Normal
Muggy	Angus	Mostly cloudy	Hot	Breezy	Alert

mapping from input to output. The rest of the rules had a degree of support less than 0·01, and were removed from the rule base.

Upon examination of Table 3, it can be shown that the set of 20 rules capture a variety of input conditions. For example, the first rule processes input data as follows: 'cattle breed of interest is Charolais, and the weather conditions are cool, cloudy, dry and windy, then respiration rate is normal.' It should be noted that the fuzzy aggregation operator is described as 'and' in the prior rule; however, in reality the model was run with a strong 'and' but allowed for a small degree of 'or' to be used in each rule.

When this model was tested against the remaining 30% of data, the model was able to account for only 27% of total variation in respiration rate (Table 1). The slope was equal to 0·31 and the intercept was 70·0. The plot of residuals appears to be mostly random, with possibly a slight positive slope [*Fig. 1*(*d*)].

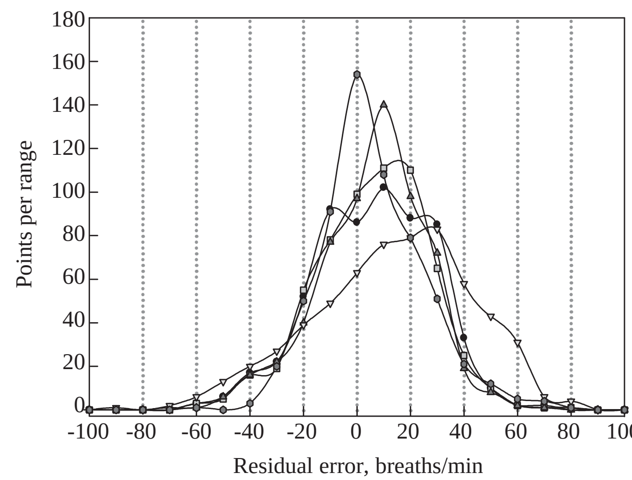

Fig. 4. Histogram of residuals for each of four models developed and tested: —⬢— linear regression; —◻— quadratic regression; —▲— data-dependent fuzzy inference system; —▽— data-free type fuzzy inference system; and —●— neural network

4.4. *Neural-network model*

The resulting neural-network model is described as having five inputs (breed, t_{db}, t_{dp}, r_s, and v_w), one hidden layer with 12 nodes, and one output (respiration rate), (5-12-1). The model parameters associated with the model are: (a) learning rate of 0·6, (b) momentum of 0·6, and (c) weights of 0·4.

When this model was tested against the remaining 30% of data, it accounted for 68% of variability associated with respiration rate (Table 1). Slope and intercept of residual plots were 0·70 and 26·3, respectively. The plot of residuals appears mostly random [*Fig. 1*(*e*)].

5. Discussion

Five models were compared side by side using two different methods. First a histogram of errors was developed counting number of errors that fell in each of 21 categories (errors between −100 and −91, −90 and −81, …, 91 and 100), for each of the four models (*Fig. 4*). In a second comparison, four sets of comparisons were made: differing breed, high and low t_{db}, high and low v_w, and high and low r_s. Approximately 20 points of similar conditions were selected from the test dataset for each condition. Table 3 shows average environmental condition, average actual respiration rate, and average result for each of the five models.

The histogram of residuals is a visual representation of distribution of errors; ideally, the plot would have a normal distribution centred at zero, with a low variance. The two regression models (linear and quadratic) and the data-dependent fuzzy inference system have a similar pattern. The only thing that differs between these models is distribution of errors between −30 and 30. The linear regression model has nearly an equal number of points in each category between −10 and 30 (approximately 90 points in each of the five categories). The quadratic regression model has almost an equal number of points in each of the categories between 0 and 20 (about 110 points in each of three categories). The data-dependent fuzzy inference system has close to a normal distribution centered at the 10 category. The data-free fuzzy inference system has an apparent bias in the histogram, with a more positive residual error. The neural network has a similar spread of error compared to the two regression models and data-dependent fuzzy inference system; however, the neural network appears to be centered at 0 unlike other models (Table 4).

It was difficult to select high and low t_{db} data points with a similar r_s, because most cool days were also cloudy days. The resulting comparison of high and low t_{db} is confounded with r_s. When evaluating responses from this comparison, heifers increased their respiration rate 50·8 breaths/min (low t_{db} of 75·0 breaths/min; high t_{db} of 125·8 breaths/min). The linear and quadratic regression models and the neural networks were able to predict a similar increase (linear regression, 56·0 breaths/min; quadratic regression, 55·0 breaths/min; neural network, 53·3 breaths/min). The data-dependent fuzzy inference system slightly under-predicted respiration rate at low t_{db} conditions and slightly over-predicted respiration rate at high t_{db} conditions;

Table 4
A subset of test data and results from five models: linear regression; quadratic regression model; data-dependent fuzzy inference system; data-free fuzzy inference system; neural network model

							Predicted respiration rate, breaths/min				
Varying parameter	*n*	*Dry-bulb temperature (t_{db}), °C*	*Solar radiation (r_s), W/m²*	*Wind speed (v_w), m/s*	*Dew point temperature (t_{dp}), °C*	*Respiration rate breaths/min*	*Linear*	*Quadratic*	*Data-dependent*	*Data-free*	*Neural network*
Dry-bulb temperature											
Low	19	22·0±1·5	435±110	13·1±1·9	16·7±0·8	75·0±17·5	72·3	75·7	72·6	82·2	76·9
High	19	34·0±0·7	762±60	13·2±2·2	18·5±1·8	125·8±25·5	128·3	130·7	136·2	108·4	130·2
Wind speed											
Low	17	30·7±1·3	849±0·3	8·4±0·3	15·1±0·3	121·5±20·0	119·1	116·0	122·4	115·0	116·5
High	17	30·6±1·0	773±54·5	22·7±4·8	15·8±2·3	80·3±19·5	95·2	86·3	77·3	92·3	83·0
Solar radiation											
Low	22	28·1±0·0	347±10·0	6·2±1·7	19·0±0·6	102·3±25·6	96·8	96·7	101·6	120·8	99·5
High	22	28·5±0·8	851±1·8	6·3±2·0	16·1±0·6	128·8±29·6	114·1	113·3	123·0	113·0	125·0
Cattle breed											
Angus	25	29·7±2·0	581±232	10·7±4·6	18·9±1·9	120·3±33·6	110·6	110·5	113·9	96·6	116·6
MARC III	25	29·4±1·7	574±223	10·3±4·0	18·6±2·2	113·5±25·7	107·5	106·4	113·7	97·1	112·4
Gelbvieh	25	29·8±2·0	572±227	8·9±3·8	18·4±2·0	98·6±25·5	105·8	106·1	105·9	125·6	100·6
Charolais	25	29·8±2·0	573±229	10·6±4·5	19·0±1·6	92·2±28·1	93·2	93·5	91·0	108·4	84·6

n, number of data.

therefore, this model slightly over-predicted the response change (63·6 breaths/min). The data-free fuzzy inference system over-predicted respiration rate at low t_{db} conditions and under-predicted respiration rate at high t_{db} conditions; therefore, this under-predicted the response change (26·2 breaths/min).

An increase in v_w from 8·4 to 22·7 caused a decrease in actual respiration rate by 41·2 breaths/min. The data-dependent fuzzy inference system was the only model that predicted a decrease (−45·1 breaths/min). The rest of the models under-predicted the change (linear regression, −23·9 breaths/min; quadratic regression, −29·7 breaths/min; data-free fuzzy inference system, −22·7 breaths/min; neural network, −33·5 breaths/min).

An increase in r_s caused an increase in respiration rate of 26·5 breaths/min. The neutral network closely predicted that change in respiration rate (25·5 breaths/min), while the data-free fuzzy inference system model incorrectly predicted a decrease in respiration rate of 7·8 breaths/min. All other models slightly under-predicted that change in respiration rate (linear regression, 17·3 breaths/min; quadratic regression, 16·6 breaths/min; data-dependent fuzzy inference system, 21·4 breaths/min).

Breeds were affected by environmental conditions differently; the effect is associated with colour of animals. Angus heifers with black hair had the highest respiration rate, then MARC III with dark red, Gelbvieh with tan, and finally Charolais with white or off-white hair. The data-free fuzzy inference system model was the only model that did not predict the correct trend. The neural-network model had the best predictions. This model closely predicts responses of Angus, MARC III, and Gelbvieh; however, slightly under-predicts Charolais response. The linear and quadratic regression models and data-dependent fuzzy inference system model all under-predict Angus, MARC III, and Gelbvieh response, while closely predicting Charolais response.

6. Conclusions

Five different models were developed to predict respiration rate, based on four weather parameters (dry bulb temperature t_{db}, dew point temperature t_{dp}, solar radiation r_s, wind speed v_w) and breed, using a 70% subset of data points and tested using the remaining 30% of data. Two regression type models were developed—one used only linear components and the other used four additional terms (significant interaction and quadratic effects). Using the test dataset, the addition of these four terms only accounted for an additional 3% of variation in the test dataset (59% compared to 62%). Two fuzzy inference systems were developed—one using the Sugeno method or data-dependent method, and the other using the Mamdani method or data-free method. It appeared that the data-dependent fuzzy inference system had similar results, maybe slightly better, than regression models. This model could account for approximately 66% of variation. The data-free fuzzy inference system development method presents a very interactive modelling environment that is more time consuming, because it relies on the modeller to describe the model in terms that are intuitive and can be understood. In this comparison, this model did not perform as well as other methods, accounting for 27% of variation in test dataset. The neural-network model accounted for approximately 68% of variation, and predicted responses slightly better than the data-dependent fuzzy inference system model. The data-dependent fuzzy inference system and the neural network both predicted respiration rate well. The data-dependent fuzzy model predicted impact of v_w better, and the neural-network model predicted impact of t_{db}, r_s, and breed better. It appeared that the neural-network model may be a slightly better approach; however, the researcher may learn more about responses using a fuzzy inference system approach. With all models tested, respiration rate is over-predicted at low-stress conditions and under-predicted at high-stress conditions; this would indicate that models are lacking a key piece of input data to make an accurate prediction; possibly the accumulative effects of prior weather conditions.

Four of the five models could be used in their current form; the data-free fuzzy inference system would require more work to make accurate predictions. Application is quite easy for all models. The regression type could easily be inserted in a spreadsheet or written in a software program. A neural network or fuzzy inference system model could be applied either using a personal computer with appropriate commercial software, or easily converted to a executable program using tools contained with in neural network/fuzzy inference system developing software. For field application, the neural network or fuzzy inference system model could be burned into a microchip for application into a controller. The advantage of a fuzzy inference system over the other types of models is the ability to input 'fuzzy' inputs; for example, the breed could be implemented as a continuous hide colour (black, dark red, red, light red, etc.) instead of one of four discreet categories.

References

Anderson K A; Magnuson B A; Tschirgi M; Smith B (1999). Determining the geographic origin of potatoes with trace metal analysis using statistical and neural network classifiers. Journal of Agricultural Food and Chemicals, **47,** 1568–1575

Babuska R (1998). Fuzzy Modeling for Control. Kluwer Academic Publishers, Norwell, MA, USA

Batchelor W D; Yang X B; Tschanz A T (1997). Development of a neural network for soybean rust epidemics. Transactions of the ASAE, **40**(1), 247–252

Belal S Y; Taktak A F G; Nevill A J; Spencer S A; Roden D; Bevan S (2002). Automatic detection of distorted plethysmogram pulses in neonates and pediatric patients using an adaptive-network-based fuzzy inference system. Artificial Intelligence Medicine, **4,** 149–165

Bezdek J C (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York, USA

Bharath R; Dorsen J (1994). Neural Network Computing. McGraw-Hill, New York, USA

Bochereau L; Bourgine P; Palagos B (1992). A method for prediction by combining data analysis and neural networks: applications to prediction of apple quality using near infrared spectra. Journal of Agricultural Engineering Research, **51,** 207–216

Brown-Brandl T M; Nienaber J A; Eigenberg R A; Mader T L; Morrow J L; Dailey J W (2003). Relative heat tolerance among cattle of different genetics. ASAE Paper No. 03-4035. St. Joseph, MI, USA

Busby D; Loy D (1996). Heat stress in feedlot cattle: producer survey results. Beef Research Report, A.S. Leaflet R1348. Iowa State University, Ames, IA, USA

Chiu S L (1994). Fuzzy model identification based on cluster estimation. Journal of Intelligence and Fuzzy Systems, **2,** 267–278

Eerikainen T; Zhu Y H; Linko P (1994). Neural networks in extrusion process identification and control. Food Control, **5**(2), 111–119

Eigenberg R A; Brown-Brandl T M; Nienaber J A; Hahn G L (2005). Dynamic response indicators of heat stress in shaded and non-shaded feedlot cattle—part 2: predictive relationships. Biosystems Engineering, doi:10.1016/j.biosystemseng.2005.02.001):

Ganjyal G M; Hanna M A; Jones D D (2003). Modelling selected properties of extruded cross-linked starches with neural networks. Journal of Food Science, **68**(4), 1384–1388

Hahn G L; Parkhurst A M; Gaughan J B (1997). Cattle respiration rate as a function of ambient temperature. ASAE Paper No. MC97-121. St. Joseph, MI, USA

Huang Y; Whittaker A D; Lacey R E (1998). Internal model control for a continuous, snack food frying process using neural networks. Transactions of the ASAE, **41**(5), 1519–1525

Inform (1999). FuzzyTech v5.3 User's Manual. Inform GmbH, Aachen, Germany

Jain A K; Dubes R C (1988). Algorithms for Clustering Data. Prentice-Hall Inc., Englewood Cliffs, NJ, USA

Jang J R (1993). ANFIS adaptive-network-based fuzzy inference system. IEEE Transaction on Systems, Man, and Cybernetics, **23**(3), 665–685

Jang J R; Sun C (1995). Neuro-fuzzy modeling and control. Proceedings IEEE, **83**(3), 378–405

Karayiannis N B; Venetsanopoulos A N (1993). Artificial Neural Networks: Learning Algorithms, Performance Evaluation, and Applications. Kluwer Academic Publishers, Norwell, MA, USA

Linko P; Uemura K; Zhu Y H; Eerikainen T (1992). Application of neural network modeling in fuzzy extrusion control. Food and Bioproducts Processing, Transactions of the Institute of Chemical Engineers, **70**(Part C), 131–137

Mamdani E H; Assilian S (1975). An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man–Machine Studies, **7**(1), 1–13

MathWorks, Inc. (2000). Fuzzy Logic Toolbox User's Guide, Version 2. The MathWorks, Inc

Pannier A R; Brand R; Jones D (2003). Fuzzy modelling of skin permeability coefficients. Pharmaceutical Research, **20**(2), 143–148

Ross T J (1995). Fuzzy Logic with Engineering Applications. McGraw-Hill, Inc., New York, USA

Sablani S S; Ramaswamy H S; Sreekanth S; Prasher S O (1997). Neural network modeling of heat transfer to liquid particle mixtures in cans subjected to end-over-end processing. Food Research International, **30**(2), 105–116

SAS (1999). User's Guide: PROC-Mixed. Ver. 8. SAS Institute, Inc Cary, NC, USA

Takagi T; Sugeno M (1985). Fuzzy Identification of Systems and its Applications to Modeling and Control. IEEE Transactions on Systems, Man, and Cybernetics, **28**(1), 116–132

Thai C N; Shewfelt R L (1991). Modelling sensory color quality of tomato and peach: neural networks and statistical regression. Transactions of the ASAE, **34**(3), 950–955

Werbos P (1974). Beyond regression: new tools for prediction and analysis in behavioral sciences. PhD Dissertation, Harvard University, Cambridge, MA, USA

Woldt W E; Dahab M F; Schulte D D; Wang Y (2002). Modelling constructed wetland wastewater treatment performance using fuzzy inference systems. ASAE Paper No. 02-3041. St. Joseph, MI, USA

Woldt W E; Dvorak B I; Dahab M F (2003). Application of fuzzy set theory to industrial pollution prevention: production system modelling and life cycle assessment. Soft Computing, **7,** 419–433

Appendix A: data-dependent fuzzy inference system

For a system with two input variables and one output variable, Sugeno type fuzzy inference system has the form (Takagi & Sugeno, 1985):

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B, \text{ then } z = f(x, y) \tag{A1}$$

where: x and y are system input variables; z is system output variable; A and B are antecedent membership functions; and $f(x, y)$ is a crisp function in the consequent. Usually $f(x, y)$ is a polynomial of the input variables x and y, but it could be any function as long as outputs produced by the fuzzy inference system can appropriately simulate the system being modelled. In this research, $f(x, y)$ was a first-order polynomial, which was expressed as

$$Z = p_1 x + q_1 y + r_1 \tag{A2}$$

Defuzzification is expressed as a weighted average Z of the consequent functions

$$Z = \frac{\sum wz}{\sum w} \tag{A3}$$

where w is the rule firing strength and z is a consequent function output.

Clustering is used to establish membership functions during the development of fuzzy inference system. Usually, cluster definition is based on spatial distances between data. In a cluster space of numerical data, the distance between any two points is less than the distance between any point in cluster and any point outside cluster (Jain & Dubes, 1988). For fuzzy clusters, the boundary between any two is fuzzy. Membership functions are established by assigning a degree of belonging for each data point to a given cluster as the fulfilment of the membership function. In this research, subtractive clustering was used to establish initial Sugeno type membership functions (Chiu, 1994; Bezdek, 1981; Babuska, 1998).

A key characteristic of artificial neural networks is their ability to learn, a process by which a neural system acquires the ability to carry out certain tasks by adjusting its internal parameters according to some learning algorithms (Karayiannis & Venetsanopoulos, 1993). The initial Sugeno type fuzzy inference system was optimised by an artificial neuro-fuzzy inference system (ANFIS) (Jang, 1993), in which the fuzzy inference system is implemented into the framework of a supervised feed-forward type artificial neural network. The training, or learning, algorithm of the artificial neural network is back-propagation, which was first introduced by Werbos (1974). The synapses in ANFIS are not given values (synaptic weights). They only indicate flow direction of data. Parameters that are adjusted are in node transfer functions. In other words, the membership function parameters are adjusted, not synaptic weights. The detailed algorithm can be found in Jang (1993) and Jang and Sun (1995).

Appendix B: data-free fuzzy inference system

The Mamdani model is a type of fuzzy relational model where each rule is represented by an IF–THEN relationship. It is also called a linguistic model, because both the antecedent and consequent are fuzzy propositions (Babuska, 1998). The model structure is manually developed and final model is neither trained nor optimised. Output from a Mamdani model is a fuzzy membership function based on rules created. Since this approach is not exclusively reliant on a dataset, with sufficient expertise on the system involved, a generalised model for effective future predictions can be obtained.

A Mamdani type model is described by Mamdani and Assilian (1975), and is expressed as follows for the same system:

$$\text{if } x \text{ is } A \text{ and } y \text{ is } B, \text{ then } z \text{ is } C \tag{B1}$$

where: x and y are system input variables; z is system output variable; A and B are antecedent membership functions; and C is a consequent membership function. Fuzzy c-Means (FCM) clustering was used to establish first approximations for Mamdani type membership functions (Chiu, 1994; Bezdek, 1981; Babuska, 1998).

The centroid method is used for defuzzification. The centroid of composed shape is computed by

$$Z = \frac{\mu_c(z) z \delta z}{\mu_c(z) \delta z} \tag{B2}$$

where z is the consequent variable and $\mu_C(z)$ is the function of the composed shape.

Even though obtaining an optimised model is not possible, the relationship can be tuned and refined. Logical relationships AND and OR are adjusted to refine the model. Refinement of the logical relationships is guided by expression of t-norms and t-conorms to adjust fuzzy relationships (Ross, 1995).

Appendix C: neural networks

Neural networks are highly sophisticated pattern recognition systems capable of learning relationships in patterns of information. They mathematically mimic the biological human learning process, and are capable of learning relationships between system inputs and responses. The mathematical relationships describing the processes may not be known (Batchelor *et al.*, 1997). Neural networks are sets of mathematical algorithms applied in processing elements (PEs) arranged to imitate the complex, non-linear and parallel mechanisms involved in interpretation of information by biological neural network.

These models learn from examples through iteration, without requiring a prior knowledge of relationships between variables under investigation (Eerikainen *et al.*, 1994; Linko *et al.*, 1992), and have performed well even with noisy, incomplete, or inconsistent data (Bochereau *et al.*, 1992).

Constructing neural-network models involves using sets of input and output vectors to train the neural network, selecting a transfer function applied to each PE, as well as starting weights applied to each interconnection between two PEs, and defining a learning rule through training. The neural network then produces its own output vectors, which are compared to the training output vectors. If desired, the degree of accuracy between neural network output and training output vectors is not satisfied, the neural-network applies the learning rule to adjust weights of interconnections, retrieves other data, and repeats comparison until the accuracy criterion is met.

A variety of learning and development strategies exist and are available for development of neural network. However, back propagation learning is the predominate method used to train neural network, as revealed in food-processing literature and other techniques, will not be discussed here.

Regardless of learning strategy, a neural network consists of a large number of highly interconnected processing elements called neurons. Each neuron receives input signals from multiple neurons in proportion to their connection weights W_{ij}. Within each neuron a threshold value (bias, B_i) is added to the weighted sum and non-linearly transformed using an activation function to generate output signals. A fully connected three-layer network is illustrated in *Fig. C1*. The response O of each neuron i to input signals I from the connecting neurons j can be mathematically expressed as

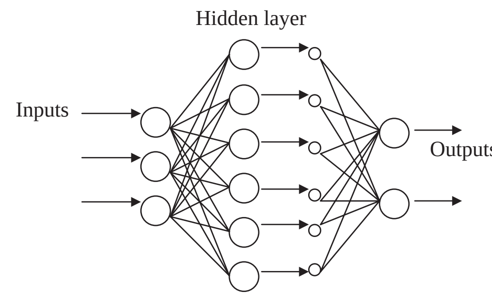

Fig. C1. A typical single layered neural network

$$O_i = f\left(\sum_{j=1}^{m} I_j W_{ij} + B_i\right) \tag{C1}$$

The transfer function f is any linear or non-linear function. Most commonly used functions are sigmoidal and hyperbolic tangent. The learning process starts with randomly initialised weights. A set of input data is presented to this network and the resulting output is compared with a corresponding desired output. Errors associated with output neurons are transmitted from output layer to input layer through hidden layers using a back-propagation algorithm, hence the name back-propagation network. In order to minimise errors, weights are adjusted at the end of each back-propagation cycle. This procedure is repeated several thousand times over the entire learning set (learning runs) (Bharath & Drosen, 1994). In principle, if a sufficient number of these input/output combinations are used for learning/training of neural network, such a trained network should be able to predict output for new inputs (Sablani *et al.*, 1997).

To prevent overfitting or overtraining, an early stopping strategy is employed to enhance the ability of networks to generalise (perform well on new data) (Anderson *et al.*, 1999). Model identification is a process of selection in which a number of models are fitted, and then the best one is determined. Determination of number of neural-network hidden nodes, along with determination of neural model order, is crucial in identification of a neural-network process model (Huang *et al.*, 1998).